

FEATURES	Source	1..883	128 Sec. 2 Yan-Chu-Yuan Road, Taipei 11529, Taiwan
gene	orgnism="Rana catesbeiana" db_xref="taxon:8400" tissue_type="liver"	1..883	
sig_peptide	/gene="rcr" 244..309 /gene="rcr" 244..645 /gene="rcr"		
cds	/note="stored in oocytes" /codon_start=1 /product="ribonuclease precursor" /protein_id="AA010702.1" /db_xref="GI:4204399" /translation="MCAKSLILVFGILGLSHSLSONMATFOOKHIIINPIINCTIT MDNNIYIVGGCKRVNFIISATVVAICGVINMVLSTPRPOLMTCRTSITPRP CPYSRTETNYICVCKENQYVHFAGIGRCP" 310..642 /gene="rcr" /function="pyrimidine-guanine sequence-specific RNase" /function="cytotoxic agent against several tumor cells" /product="ribonuclease" 827..832 /gene="rcr"		
BASE COUNT	285 a 186 c 150 g 262 t		
ORIGIN			
Query Match	57.9%; Score 191.2; DB 5; Length 883; Best Local Similarity 75.6%; Pred. No. 1.2e-44; Matches 251; Conservative 0; Mismatches 78; Indels 3; Gaps 1;		
1	cagaactggagctactcttcacgcagagaacatatcatcaactccgaatc--ctgcagac 57		
310	CAGAACTGGGCAACATTTTCAGCAGAGACACATTTAAACACACCGCATCATCACTGTAAAC 369		
58	actatcatgagacaacacatctacatcgltgtygtgcagtgcaaacgctttaacacttcc 117		
370	ACCATCATGGACAAACAAATATATATATTCGTAGAGAGGTCAATGCAAGAGATGAACACTTTC 429		
118	atcacctcttcctgcctacactggttaaacctatctgcactgtygtatataacatgaagtt 177		
430	ATAATTTCTTCGCAACACCGGTGAAGCCACTGTGTACCGGGGTGATTAATATGAAATGA 489		
178	ctgtctactactcgtttccagctggaacacttgcactgcgttaacttatactaccgcgttcg 237		
490	TTAACTACCAACAAGATTCACAGCTCAACACTTGTCATCTGATTCATTAACAACCCGCCCT 549		
238	tgcgcgtactcttcttcgtaactggaactatacatctcgtttaaatgcgaaaacagatgac 297		
550	TGTCATATTAAGCTCCAGAACCGGAAACTAATTAATGTAATAATGTGAGATCAATATAT 609		
298	ccggtcatttcgcgtgcatcgtcgttcgcc 329		
610	CCCGTACATTTTGTCTGGAATAGGACGATGTCC 641		
RESULT 2			
AF351209	402 bp DNA linear VRT 29-JUN-2001		
LOCUS			
DEFINITION	Rana catesbeiana RNase A-type ribonuclease rc208 precursor, gene,		
ACCESSION	AF351209		
VERSION	AF351209.1		
KEYWORDS	GI:13560895		
SOURCE	bullfrog.		
ORGANISM	Rana catesbeiana		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.		
REFERENCE	1 (bases 1 to 402)		

FEATURES	source
AUTHORS	Rosenberg, H. F., Zhang, J., Liao, Y. D. and Dyer, K. D.
TITLE	Rapid diversification of RNase A superfamily ribonucleases from the
JOURNAL	bullfrogs, Rana catesbeiana
MEDLINE	J. Mol. Evol. 53 (1), 51-58 (2001)
PUBMED	11683320
REFERENCE	2 (bases 1 to 402)
AUTHORS	Rosenberg, H. F., Zhang, J., Liao, Y. D. and Dyer, K. D.
TITLE	Direct Submission
JOURNAL	Submitted (21-FEB-2001) Laboratory of Host Defenses, NIAID, National
	Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA
	Location/Qualifiers
	1. .402
	/organism="Rana catesbeiana"
	/db_xref="taxon:8400"
sig_peptide	1. .66
mrna	<1. .>402
	/product="RNase A-type ribonuclease rc208 precursor"
CDS	1. .402
	/codon_start=1
	/product="RNase A-type ribonuclease rc208 precursor"
	/protein_id="AAK30255.1"
	/db_xref="GI:13560896"
	/translation="WCASKDLVFGILGLSHSLSONMAFPQKHNTNTSSINCNIT
	MDNNYITVGGCKGVNTFIISATVTAICTGYINNNVSTTRPOLNCTRTITPRP
	CPYSSRTENNYICVCKENQYVPHRAGAGCP"
BASE COUNT	120 a 93 c 75 g 114 t
ORIGIN	
Query Match	56.0%; Score 184.8; DB 5; Length 402;
Best Local Similarity	74.4%; Pred. No. 8.1e-43;
Matches 247; Conservative	0; Mismatches 82; Indels 3; Gaps 1;
QY	1 cagaactgggctactctccagcaagaatatcatcacactccg--atcatctgcac 57
DB	67 CAGAACTGGGCACACTTTCAGCAGACACATTAACAAATACATCGAGATCAACTGTAC 126
QY	58 actatcatggaacaacaacatcatcgtctggtygctagcgcaaacgyltaaaccttc 117
DB	127 ACCATCATGAGCAACAAATATATATTCGTGGGAGGTCAATCAAGGAGTGAAACACTTTC 186
QY	118 atcatctcttctgctactacacggttaagcttcttcgacttggttataataagaagtt 177
DB	187 ATAAATTTCTTTCGACACACCGTGAAGGCCATCTGTACCGGGGTGAATAATGAATGA 246
QY	178 ctgtctactactcgtttccagctggaacactctgcactcgttactctatcatcctcgctccg 237
DB	247 TTAAAGTACCACAAAGATTCCAGCTCAACACTTGCACTGCTACTTGTATTAACACCCGGCCT 306
QY	238 tgcgcgtactctcttcgctactggaactactacatctgycgtltaaatgcgaaccgagtc 297
DB	307 TGTCCATTTAGCTCCAGAACGGAAAAATATTAACAATGTGTGAATAATGGAATAATATAT 366
QY	298 ccggttcatttcgctgctgctatcgctgctgctcc 329
DB	367 CCTGTACATTTTGTCTGGAATAGGACGATGTCC 398
RESULT 3	
AF351210	402 bp DNA linear VRT 29-JUN-2001
LOCUS	Rana catesbeiana RNase A-type ribonuclease rc212 precursor, gene,
DEFINITION	complete cds.
ACCESSION	AF351210
VERSION	AF351210.1 GI:13560897
KEYWORDS	
SOURCE	bullfrog.
ORGANISM	Rana catesbeiana
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana;
REFERENCE	1 (bases 1 to 402)
AUTHORS	Rosenberg, H. F., Zhang, J., Liao, Y. D. and Dyer, K. D.

TITLE	rapid diversification of RNase A superfamily ribonucleases from the bullfrog, <i>Rana catesbeiana</i>
JOURNAL	J. Mol. Evol. 53 (1), 31-38 (2001)
MEDLINE	21539506
PUBMED	11683320
REFERENCE	2 (bases 1 to 402)
AUTHORS	Rosenberg, H.F., Zhang, J., Liao, Y.-D. and Dyer, K.D.
TITLE	Direct Submission
JOURNAL	Submitted (21-FEB-2001) Laboratory of Host Defenses, NAID, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES	Location/Qualifiers
SOURCE	1..402
	/organism="Rana catesbeiana"
	/db_xref="taxon:8400"
slg_peptide	1..66
MRNA	<1..>402
	/product="RNase A-type ribonuclease rc212 precursor"
CDS	1..402
	/codon_start=1
	/product="RNase A-type ribonuclease rc212 precursor"
	/protein_id="AAK30256.1"
	/db_xref="GI:13560898"
	/translation="MCAKSPILVVGILILVLSHLSLSONMNAFTPOKHITNTSSINCSNI MNSSLYIVGGQCKKRVKFTFSSATFVSGVTDKRVLSSTKFDLIRITFTPPRP CPYSTRETNITVCKCENQYVPHAGVGQCP"
BASE COUNT	120 a 94 c 76 g 112 t
ORIGIN	
Query Match	47.8%; Score 157.6; DB 5; Length 402;
Best Local Similarity	69.3%; Pred. No. 5.9e-35;
Matches 230; Conservative	0; Mismatches 99; Indels 3; Gaps 1;
OY	1 cagaactggtgactcttcagcagaacaatcatcaactcgc--atcatctgcac 57
Db	67 CAGAACTGGGCAACCTTTCAGCAGACACATTAACAAATACATCGACATCAACTGTAGC 126
OY	58 actaaltygaacaacaatctacatcgttggtyggtcagtgcaaacgtttaacattc 117
Db	127 AACATCATGAAACAACAGCTTATATATCGTGGAGAGTCAATGCAAGAAAGTGAACTTTC 186
OY	118 atcatctctctgtctactactcgttaagaatctcgcactcgtgtgtatataacatgaagtt 177
Db	187 ATAGCTTCTTGTGCACACCGTGAAGGACATCTCTTACCGGGGATMACCATTAAGAAAGTT 246
OY	178 ctgtctactactcgtttccagctgaaacactcgcactcgtactctatacccgsgtgc 237
Db	247 TTAACTATACAAAAATTCACAGTCGACATTTTGACATCGTAATTTTCATTACACCCCGCCT 306
OY	238 tgcgcgactctctcgtactcgaactactactactcgtttaaatggaaccagatc 297
Db	307 TGTCATATATAGCTCCAGAAAGAACTAATTAACATATGTGTAATAATGTGAGATCAATATAT 366
OY	298 ccggttcattcgcgtgtatcgttcggttcgcc 329
Db	367 CCCGTACATTTTGTGTGAATAGACAAATGTCC 398
RESULT 4	
AF351207	
LOCUS	AF351207 399 bp DNA linear VRT 29-JUN-2001
DEFINITION	Rana catesbeiana RNase A-type ribonuclease rc203 precursor, gene, complete cds.
ACCESSION	AF351207
VERSION	AF351207.1 GI:13560891
KEYWORDS	
SOURCE	bullfrog.
ORGANISM	Rana catesbeiana
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
REFERENCE	1 (bases 1 to 399)
AUTHORS	Rosenberg, H.F., Zhang, J., Liao, Y.D. and Dyer, K.D.
TITLE	Rapid diversification of RNase A superfamily ribonucleases from the

FEATURES	source	location/Qualifiers
JOURNAL	bullfrog, Rana catesbeiana	
MEDLINE	J. Mol. Evol. 53 (1), 31-38 (2001)	
PUBLISHED	21539506	
REFERENCE	11683320	
AUTHORS	2 (bases 1 to 399)	
TITLE	Rosenberg, H.F., Zhang, J., Liao, Y.-D. and Dyer, K.D.	
JOURNAL	Direct Submission	
FEATURES	Submitted (21-FEB-2001) Laboratory of Host Defenses, NAID, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA	
source	1. 399	
sig_peptide	/organism="Rana catesbeiana"	
mRNA	/db_xref="taxon:8400"	
CDS	1. .66	
	<1. .>399	
	/product="RNase A-type ribonuclease rc203 precursor"	
	1. 399	
	/codon_start=1	
	/product="RNase A-type ribonuclease rc203 precursor"	
	/protein_id="AAK30253.1"	
	/db_xref="GI:13560892"	
	/translation="MCAKSLILVGIILGSLISLSOMWAKPOOKHIPSINICNTITMDNNIYVGGCKRVNFIPISSATVVAIAICGVYNSNVLSPTKQLDTCRTSITPPPCPSYKRETNKICVCKENQLEHFDAGK"	
BASE COUNT	123 a 94 c 73 g 109 t	
ORIGIN		
Query Match	47.5%; Score 156.6; DB 5; Length 399;	
Best Local Similarity	70.0%; Pred. No. 1.2e-34;	
Matches 226; Conservative 0; Mismatches 94; Indels 3; Gaps 1;		
QY	1 cgaagctgggactcttccagcagaagaacatcatcacaacacgcg---atcatctgcaac 57	
DB	67 CAGACTGGGCGCAAATTTTCAGCAGACGACATCTCAAGACACATTCAGCATCAACTGTATAC 126	
QY	58 actatcatggaacaacaacatctacatcglttggtgagtcagtgcaaacglttaacacttic 117	
DB	127 ACCATCATGTGACAAACAATATATATATCGTGGAGGTCATGCAAGAAAGTGAAACCTTTC 186	
QY	118 atcatctcttctgctactactcgtttaagcatctgcagctggtgtatataacatgaagctt 177	
DB	187 ATAAATTTCTTCGACACACACCGTGAAGGCGCATCTTAACGGGGTGACAAATAGTAAATGTA 246	
QY	178 cgtctactaacccgtttccagctggaacatctgcagctgactctatcatctccggtctc 237	
DB	247 TTAAGTCCACACAGATTCCAGCTGCACACTTCGACTTCGATTCATTCACCCCGGCGCT 306	
QY	238 tgcgcgtactctctctcgctactgaaactaactacatctcgtttaaatgcaaacacagtac 297	
DB	307 TGTCCATATAGTGTCCAAAAGGAACCTAATAGATATGTGTAATGTGAGAGATCAACTT 366	
QY	298 ccggttcatttcgctggtatcg 320	
DB	367 CCCGTACATTTTGTGTAATAGG 389	
RESULT 5		
AF351211		
LOCUS	AF351211 402 bp DNA linear VRT 29-JUN-2001	
DEFINITION	Rana catesbeiana RNase A-type ribonuclease rc218 precursor, gene, complete cds.	
ACCESSION	AF351211	
VERSION	AF351211.1 GI:13560899	
KEYWORDS		
SOURCE	bullfrog.	
ORGANISM	Rana catesbeiana	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.	
AUTHORS	1 (bases 1 to 402)	
TITLE	Rosenberg, H.F., Zhang, J., Liao, Y. D. and Dyer, K.D.	
	Rapid diversification of RNase A superfamily ribonucleases from the bullfrog, Rana catesbeiana	

REFERENCE	2 (bases 1 to 884)
AUTHORS	Liao Y.-D., Huang, H.-C., Leu Y.-J., Wei, C.-W., Tang, P.-C. and Wang, S.-C.
TITLE	Direct Submission
JOURNAL	Submitted (18-JUL-2000) Institute of Biomedical Sciences, Academia Sinica, No.128 Sec.2 Yan-Chu-Yuan Road, Taipei 11529, Taiwan 3 (bases 1 to 884)
REFERENCE	Liao Y.-D., Huang, H.-C., Leu Y.-J., Wei, C.-W., Tang, P.-C. and Wang, S.-C.
AUTHORS	Direct Submission
TITLE	Submitted (08-AUG-2001) Institute of Biomedical Sciences, Academia Sinica, No.128 Sec.2 Yan-Chu-Yuan Road, Taipei 11529, Taiwan
JOURNAL	Sequence update by submitter
REMARK	On Aug 8, 2001 this sequence version replaced gi:1118657.
COMMENT	Location/Qualifiers
FEATURES	<p>source</p> <p>1..884</p> <p>/organism="Rana catesbeiana"</p> <p>/db_xref="taxon:8400"</p> <p>/tissue_type="liver"</p> <p>1..685</p> <p>/product="RC-RNaseL1 ribonuclease precursor"</p> <p>1..219</p> <p>220..282</p> <p>220..618</p> <p>/codon_start=1</p> <p>/product="RC-RNaseL1 ribonuclease precursor"</p> <p>/protein_id="AAG30414.2"</p> <p>/db_xref="GI:15111755"</p> <p>/translation="MCAKSLVFGFLGLSLSSONMAKFEKHITSSIDCNTITDKAIYVIGKCKERNFTLISSEDNVAKISGVSPDRKELSTPSKLNATCINDSTIPRCDKHPSPDNKKICKCKKQLPVHVGIGKC"</p> <p>283..615</p> <p>/product="RC-RNaseL1 ribonuclease"</p> <p>619..884</p>
BASE COUNT	273 a 199 c 154 g 258 t
ORIGIN	
Query Match	36.2%; Score 119.6; DB 5; Length 884;
Best Local Similarity	64.7%; Pred. No. 6.1e-24;
Matches 211; Conservative 0; Mismatches 109; Indels 6; Gaps 2;	
QY	1 cagaactggctactcttcacagcagaataatcatcaacactcgcatac---tgcaac 57
Db	283 CAGAACTGGCGAAATTTAAGAGAGACACATTACACACATCGACATCATTTGTAAC 342
QY	58 actatcatgagcaacaacatactacatcgctgtgtgctgagtaaacggttaacattc 117
Db	343 ACTATCATGAGCAAAAGCCATATATATCGTAGAGGAAATGAAAGAACGGAACTTTC 402
QY	118 atcatctctctgtctactacactcgttaagctactctgcactggtgt---tatcaacatgaac 174
Db	403 ATCATTTCTTCTGAAGACAGCTGAAGGCCATCTGTAGCGGGGTGTCAACCCGATAGGAAG 462
QY	175 gtctgttactactcgttttccagctgtaacactctgcactcgtactctatcaactccgagt 234
Db	463 GAATTAAGTACACACAAAGTTTCAAACTCAACTTCATTCGTGATTCATTCATTCACCCCGC 522
QY	235 ccggtgcgcgtactctctcctgcgtactgaaactactatcatctcgtgttaaatgcaaacag 294
Db	523 CCTGTCTCATATCAACCCGAGCCGGATATATATAAGATATGTGTAATAAGTGAAGAACAA 582
QY	295 taccoggtcatcttcgcgtgatcgg 320
Db	583 CTTCCTCCGTACATTTGTGTGGAATAGG 608
RESULT	8
AC005674	AC005674 199450 bp DNA linear PRI 27-FEB-2001
LOCUS	AC005674 Homo sapiens chromosome 4 clone RP11-448G15, complete sequence.
DEFINITION	AC005674
ACCESSION	AC005674.11 GI:13129492
VERSION	

	KEYWORDS	HTG.
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	1 (bases 1 to 199450)	
JOURNAL	DOE Joint Genome Institute and Stanford Human Genome Center.	
REFERENCE	Direct Submission	
AUTHORS	Unpublished	
TITLE	2 (bases 1 to 199450)	
JOURNAL	Stone,N.E., Schmitz,J.J., Cox,D.R. and Myers,R.M.	
REFERENCE	Submitted (11-FEB-1998) Department of Genetics, Stanford Human	
AUTHORS	Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA	
TITLE	3 (bases 1 to 199450)	
JOURNAL	DOE Joint Genome Institute and Stanford Human Genome Center.	
COMMENT	Direct Submission	
	Submitted (27-FEB-2001) DOE Joint Genome Institute, 2800 Mitchell	
	Drive, Walnut Creek, CA 94598, USA	
	On Feb 27, 2001 this sequence version replaced gi:5757495.	
	Finishing completed at Stanford Human Genome Center	
	www.shgc.stanford.edu	
	Quality: Phrap Quality >=40 99.9% of Sequence;	
	Estimated Total Number of Errors is 0.2.	
	STS Content:	
	SHGC-S1391 G34042	
	SHGC-S69010 G41975	
	SHGC-S50921 G33945	
	SHGC-S25155 G31610	
	SHGC-S15591 G34129	
	SHGC4-1311 G33196.	
FEATURES	Location/Qualifiers	
source	1..199450	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="4"	
	/clone="RP11-448615"	
BASE COUNT	54835 a 46671 c 45223 g 52721 t	
ORIGIN		
	Query Match 13.8%; Score 45.4; DB 9; Length 199450;	
	Best Local Similarity 48.3%; Prid. No. 0.028;	
	Matches 127; Conservative 0; Mismatches 136; Indels 0; Gaps 0;	
Oy	18 ccagcagaacatcatcatcaaacctcgatcatgcaaacactatcatgaagaacaacat 77	
Db	95529 CCACCATATCAACATCACGCCACCACCCAGCATTCTACCCACATCGCCACCATTCAC 95470	
Oy	78 ctacatcgttgttgtagtgcgaacagtgttaaacaccttcacatctcttcgtctaac 137	
Db	95469 TATCATTAACCACTTACTACTACCAACCAACCACTATCAACATCATCATTTGGCAACAACAT 95410	
Oy	138 ttgtaaagctatctgcactggtgtataaacaatgaacglttcgttactactcgttcca 197	
Db	95409 CATCAACATCTATCAACCAACCAACCATCATCATCAATCATGACCAATCATACCAACAC 95350	
Oy	198 gctgaacactgcgaactggtactctatatactccggtcgggtcgcgtactctcttcgtaac 257	
Db	95349 CACCATCTACTACCATTTACCAACCAACCATCATCAACGTCACATCGGCATCATTTACATGAC 95290	
Oy	258 tgaactaatcaatcgtcgatta 280	
Db	95289 TACCACTACCAACATCATCATTTA 95267	
RESULT	9	
LOCUS	HZ067264 14235 bp DNA linear VRL 31-JAN-2001	
DEFINITION	Helicoverpa zea nuclear polyhedrosis virus, partial sequence.	
ACCESSION	U67264 Z12117 U03418	
VERSION	U67264.1 GI:2078305	
KEYWORDS		

SOURCE	ORGANISM
REFERENCE AUTHORS TITLE	Helicoverpa zea single nucleocapsid nucleopolyhedrovirus. Helicoverpa zea single nucleocapsid nucleopolyhedrovirus viruses; dsDNA viruses, no RNA stage; Baculoviridae; Nucleopolyhedrovirus. 1 (bases 335 to 1075) Cowan,P., Bulach,D., Goode,K., Robertson,A. and Tibbe,D.E.
JOURNAL MEDLINE REFERENCE AUTHORS	Nucleotide sequence of the polyhedrin gene region of Helicoverpa zea single nucleocapsid nuclear polyhedrosis virus: Placement of the virus in lepidopteran nuclear polyhedrosis virus group II J. Gen. Virol. 75 (Pt 11), 3211-3218 (1994) 95053907 2 (bases 1 to 14235) Le,T.H., Wu,T., Robertson,A., Bulach,D., Cowan,P., Goode,K. and Tibbe,D.
TITLE	Genetically variable triplet repeats in a RING-finger ORF of Helicoverpa species baculoviruses
JOURNAL MEDLINE PUBMED REFERENCES	Virus Res. 49 (1), 67-77 (1997) 91784198 3 (bases 1 to 14235) Le,T.H., Wu,T., Robertson,A., Bulach,D., Cowan,P., Goode,K. and Tibbe,D.
TITLE	Direct Submission
JOURNAL	Submitted (20-AUG-1996) Microbiology, U. of Melbourne, Grattan Street, Parkville, VIC 3052, Australia
COMMENT FEATURES SOURCE	On May 14, 1997 this sequence version replaced gi:60442 gi:454319. Location/Qualifiers 1..14235 /organism="Helicoverpa zea single nucleocapsid nucleopolyhedrovirus" /strain="Elkar (Elcar)" /db_xref="taxon:10468" 335..1075 /codon_start=1 /product="polyhedrin" /protein_id="AAB54089.1" /db_xref="GI:2078306" /translaton="MYYRYSYSPITGKTYVYNKKYFKNGIKAVIKNAKRKHLEHEHE ERNLSDSKQLVAEDPLGCGKNQKLTFEISVKPKDMLVMWGSREPLEETWR FMESFPFINDOEFLIMDFELSYNMHPKPNRCGRALADHALRCDDYTPHYIVIRVES YVGNSPFYRISSIAKKYGCGPMVNHAYTNSFEFITNVAMENKYRIYVGVDSAE ELLIELSLFKIFEPAPDAPIYGPAY" complement(1072..2310) /note="AcMNPV ORF8/ORF1629 homolog" /codon_start=1 /protein_id="AAB54090.1" /db_xref="GI:2078307" /translaton="MVOLQSVOYLAKKNTNTINNVDPLLRMISAPSLKRTYSANAQ SNSRQLNWYDCIQLTKLAQEYINNOVLIDIGRSIIIPLSOTMSVPVPPKTTIPK PPPVSTSDIMOKTISPRIIIDSSESSQOTLMPEIPPPPPOGTNLSSTIPPP PPPPSPLEILEDIVISSKDREDYTSTRVOLKDPRELMEQIQKGILKLVKSPPDG SIYTVTAASPTAKILORRNAVQSVSESSESGMTDEOQOBSASELKOVRSLY NITLWSMIKNYSISTEAODTLTISKLNORLNSNOTOOISAKQIFEDNLIOQT DNPLDKDNVELGLSDKNLQQFIHAYEDLTFFKKDYDALMWASLGATKPTNKLDEL RSNLDTKMTSKLTMSTESQV" 2325..3128 /gene="pk1" 2325..3128 /gene="pk1" /codon_start=1 /protein_id="AAB54091.1" /db_xref="GI:2078308" /translaton="MDDRFYKEINQFAEKTIQNVRALVDGKFGKMCVIKHEPTGKLF VKRSVAIKRYTEIEPLVHOLMKORRYITIKYLSLTQLSOILLDYVAGGDIDPELKK HKRVSAETSINSLOLTBALNALHSYIIINDLKEVLIVRHQIITYCYGCKLYTN TSSCRDOTEKYMSEPCLKRONVDYHVMAALGITFELLIGHPYKSNONNEEDFDLD VLQROORKRLHKRYFSSDAQKFLAEMALMNINRYCTGYEVYIKHPSL" 3105..3164 /note="dyad element" 3178..3203 /rpt_type-direct /rpt_type-direct 3224..3249 /rpt_type-direct

[illegible]

CDS	complement(6791..6961) /note="orf7" /codon_start=1 /product="unknown" /protein_id="AA053750.1" /db_xref="GI:12483758" /translation="MYQFFRTLVACEPLTKINSIIINLSITYIYSAGYQCQCHMYVC ALICHEKGLNCY" 6949..7808
CDS	/note="orf8" /codon_start=1 /product="1e-0" /protein_id="AA053751.1" /db_xref="GI:12483759" /translation="MSGTILKILYLDIDSDSDAKLEPRYSEMPSPASQOMNNAVDE IDVAVIKCFILKNMSSDVTYTNARQNVNKLAAVLIVLDEKQYKKNLKHSHLYLK ETSSVYITLDEDOCHHTLPLIQDLTLITYLNMFSDEVNYSQKFIPLPYLKLK ILKLFQYDCKCALTKQLQALNTLLTOSADSKKHAIINOSQVILVPLENPLYECN ICRPTFNDERHIKRNPECCGYKICMLCYANILMKYSTVPTCPVCSTKSSSVSEKQV YTAFTDNI" 7823..9229
CDS	/note="orf9" /codon_start=1 /product="p49" /protein_id="AA053752.1" /db_xref="GI:12483760" /translation="MNLDEKNVALERNNYKYLFLASYNLADQGLSTSSKPIREVL YNNNNINDDASLSLGLDYLDLIGLNNYILDRDVMKTYIKPQFFVCTKKVELKRD ORVYIKFDPIVATINFEVKNPSEKFLINFESSVIDKRNFNVDKNYCLIQNTGTGM FDQAVDMCGVRMCEVPRIELESSPEPYRLYLVGDDAMARHATNNISFDSGNFLKNP YKGLPMERTNYKIIINSKFPPTKKNPHLNEFKQEDYKSAVVKIQDYLIDYDAKAPD DLKDLNEMHTYISVYKFEVTKEMDEGEPGNYSEIYIDRAVDKQKLSIKIDETTM EPFLRYDPSYIFIRPDILQIKGLTMNFYVPHKFLFALINNSLGSSTLLEPRKIL PYRQYQPPYRIINDETYVVDKKIKYILKYTFPANTIPAYLILRDGYESSSEIKITLIDLK PMVQNTILKLLILNAPPSK" 9240..9485
CDS	/note="orf10" /codon_start=1 /product="odv-e18" /protein_id="AA053753.1" /db_xref="GI:12483761" /translation="MDLKRGTYYTGAGRPANMNLNPSMLTILIALYIIILLMLFOS SSPESKAGDTNAPFQNPFLNATMRNPFVNTPORTMM" 9500..10354
CDS	/note="orf11" /codon_start=1 /product="odv-e27" /protein_id="AA053754.1" /db_xref="GI:12483762" /translation="MKRFKCSQNKIRTVTEIINADEKLHKDYDLADFNANKNLNLSLEY DNLQIKILAKYMLMLRELTOPLTLFRDKNARIIVSIVFASLGFVINRVNPMJIN HFNKSMFEIYVENNNASIPGEILRFECOHGNDVYCIDRPSIOLMSLDEPDLVDVYNN MHKRNKYMIKATPCRCAPKRHSREREPPLIEILVLEDTVOYMTLFIENHAYLHY ILKNYGVADYSRSLSDHTLFSNKSRLPNMKFSMLLSKRFSEIDYDSLINTKNTNKN IGLITYND" 10399..10677
Query Match	13.7%: Score 45.2; DB 14; Length 131403;
Best Local Similarity	52.1%: Pred. No. 0.03;
Matches 101; Conservative	0; Mismatches 93; Indels 0; Gaps 0;

Db	4186	AGTATCTTCATTGACGATGCATGTGCACCAACTTACTTATTTGTCGTCATCATCATTAC	4245
OY	211	actgcgtacctctat 224 	
Db	4246	ATTAGTAGTACTATTAT 4259	
RESULT 12			
AC095975/c			
LOCUS			
DEFINITION	AC095975	92536 bp DNA	linear HTG 20-DEC-2001
ACCESSION	Rattus norvegicus clone CH230-ILL19,	*** SEQUENCING IN PROGRESS	
VERSION	AC095975.2	GI:17943615	
KEYWORDS	HTG: HTGS PHASEL.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1. (bases 1 to 92536) Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Aisbrook,S.L., Amaralunge,H.C., Aye,J.R., Baks,T., Barbara,J., Benton,U., Blimey,K., Blankenburg,K., Bonini,D., Bouck,J., Bowles,S., Byleva,M., Brown,E., Brown,M., Bryant,N.P., Buhaey,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F., Cartier,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dean,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,U., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homel,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolyvet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C., Kravtovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Melker,M., Miner,G., Miner-Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nnokenkwo,S., Ogulu,M., Okunnu,G., Oragunye,N., Oyedero,R., Pace-A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primm,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojudoakan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Stilson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabot,P., Telford,B., Thomas,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Tamirad,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalona,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlaczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G., and Gibbs,R.		
TITLE	Unpublished		
JOURNAL	Direct Submission		
REFERENCE	2 (bases 1 to 92536)		
AUTHORS	Worley,K.C.		
JOURNAL	Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
COMMENT	On Dec 20, 2001 this sequence version replaced gi:15627595. ----- Genome Center Center code: BCM		

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
13	AC084361/c	172853 bp	DNA	linear	PRI 27-SEP-2001				
		Human sapiens 12q BAC RP11-364C11 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	AC084361						
		AC084361.19 GI:15027716							
		human.							
		Homo sapiens							
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
		1 (bases 1 to 172853)							
		Muzny D.M., Adams C., Adio-Ondola B., Alt-osman F.R., Allen C., Alsbrooks S.L., Amaraltinge H.C., Are J.R., Banks T., Barbara U., Benton J., Blmage K., Blankenburg K., Bonnin D., Bouck J., Bowie S., Brileva M., Brown E., Brown M., Bryant N.P., Bunay C., Butch P., Burkett C., Burrell K.L., Byrd N.C., Caron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R., Chen Z., Chiu D., Chowdhry I., Christopoulos C., Cleveland C.D., Cox C., Coyle M.D., Dalhorne S.R., David R., Davila M.L., Davis C., Day-Carroll L., Dederich D.A., Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H., Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J., Earhart C., Edgar D., Edwards C.C., Elhaj C., Emerling S., Escotto M., Falls T., Ferriguo D., Flagg N., Ford J., Foster P., Frintz P., Gebisi A., Gao J., Garcia A., Garner T., Garza N., Gill R., Gottell J.H., Guvera W., Gunaratne P., Hale S., Hamilton K., Han J., Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J., Hernandez O., Hodgson A., Hogues M., Holloway C., Hollins B., Homsi F., Howard S., Huber J., Hulik S., Hume J., Ioshikhes I., Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolivet S., Joudan S., Karlsson E., Kelly S., Khan U., King L., Korvan J., Kovar C., Kratovic J., Kureshi A., Landry N., Leal B., Lee E., Lewis L.C., Lewis L., Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W., Louisged H., Lozado R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J., Maheshwari M., Mapua P., Marondel I., Martin R., Mattindale A., Martinez E., Massey E., Mawhney E., McLeod M.P., Meador M., Mel G., Mercher S., Metzker M., Miller A., Miner G., Miner Z., Mitchell T., Monabbat K., Montgomery K.T., Morgan M., Morris S., Moser M., Neal D., Nelson D., Newton J., Newton N., Nguyen A., Nguyen N., Nguyen N., Nickerson E., Nwokenwo S., Ogih M., Okumou G., Oragunye N., Oyedelo R., Pace A., Payton B., Peery J., Perez L., Peters L., Pickens R., Primus E., Pu L.L., Qulles M., Ren Y., Rives M., Rojs A., Rojodokan I., Rolle M., Ruiz S., Savary G., Scherer S., Scott G., Shen H., Shm C., Shooshitari N., Sisson I., Sodergren E., Sonalike T., Sparks A., Stanley R., Stone H., Sutton A., Svatek A., Tabore P., Tamerisa A., Tamerisa K., Thoms S., Usmani K., Vasquez L., Varro V., Villalona D., Vinson R., Wall R., Wang S., Ward-Moore S., Warren R., Washington C., Wallington S., Williams G., Williamson A., Wleczek R., Wooden S., Worley K., Wu C., Wu Y., Wu Y.F., Zhou J., Zorilla S., Kucheraipati R., Weinstein G. and Gibbs R.							
		Direct Submission							
		Unpublished							
		2 (bases 1 to 172853)							
		Worley K.C.							
		Direct Submission							
		Submitted (26-OCT-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA							
		3 (bases 1 to 172853)							
		Worley K.C.							
		Direct Submission							
		Submitted (28-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA							
		4 (bases 1 to 172853)							
		Worley K.C.							
		Direct Submission							
		Submitted (27-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One							

COMMENT

Baylor Plaza, Houston, TX 77030, USA
On Jul 28, 2001 this sequence version replaced g1:14851680
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

unpublished.) for Human and Mouse sequences

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-4002) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8080/quality.info/genbank.annotation.html>.

Source

QUALSTAT-REPORT.

misc_feature

region complement(2. .154)

repeat_region

```
/rpt_family="MIR"
```

repeat_region

```

/rpt_family="(A)n"

```

region	complement(1972..2046)
--------	------------------------

repeat_region

```
/rpt_family="L2"
```

-region	3474.	.3525
---------	-------	-------

repeat_region

```
/rpt_family="MLT1I"
```

•

region	5059.	.5355
--------	-------	-------

repeat_region

```
/rpt_family="MIR"
```

```

repeat_region      5578..5598
                    /rpl_family="AT_rich"
repeat_region      complement(5697..5795)
                    /rpl_family="L2"
repeat_region      complement(6215..6472)
                    /rpl_family="MIR"
repeat_region      7183..7560
                    /rpl_family="MER69"
repeat_region      complement(8700..8878)
                    /rpl_family="MER5B"
repeat_region      8917..9111
                    /rpl_family="MIR"
repeat_region      9212..9324
                    /rpl_family="MIR"
repeat_region      10745..10794
                    /rpl_family="CT-rich"
repeat_region      complement(10981..11063)
                    /rpl_family="MIR"
repeat_region      11214..11412
                    /rpl_family="MIR"
repeat_region      11474..11741
                    /rpl_family="AluYo"
repeat_region      12403..12453
                    /rpl_family="(CA)n"
repeat_region      12879..13046
                    /rpl_family="MIR"
repeat_region      complement(13049..13119)
                    /rpl_family="L2"
repeat_region      13376..13412
                    /rpl_family="L2"
repeat_region      complement(13414..13706)
                    /rpl_family="AluSx"
repeat_region      13854..14213
                    /rpl_family="L1ME"
repeat_region      complement(14639..14859)
                    /rpl_family="MER20"
repeat_region      14978..14999
                    /rpl_family="AT_rich"
repeat_region      15220..15387
                    /rpl_family="MIR"
repeat_region      complement(16602..16719)
                    /rpl_family="MIR"
repeat_region      complement(17029..17225)
                    /rpl_family="L1ME"
repeat_region      complement(17402..17591)
                    /rpl_family="L1ME"
repeat_region      18302..18502
                    /rpl_family="L1ME3A"
repeat_region      18542..18863
                    /rpl_family="L2"
repeat_region      complement(18973..19096)
                    /rpl_family="L1P1"
repeat_region      19098..19244
                    /rpl_family="L2"
repeat_region      19300..19687
                    /rpl_family="MLT1C"

Query Match      13.1% Score 43.2; DB 9; Length 172853;
Best Local Similarity 56.2% Pred. No. 0.12; Mismatches 63; Indels 0; Gaps 0;
Matches 81; Conservative 0;

```

RESULT 14 AC073376/c LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM

```

AC073376      178150 bp DNA linear HTG 28-JUN-2000
Homo sapiens clone RP11-173C19, WORKING DRAFT SEQUENCE, 37
unordereed pieces.
AC073376      GI:8780662
HTG: HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 178150)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-173C19
Unpublished
2 (bases 1 to 178150)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Lander,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margis,N.,
McCarthy,M., McKean,P., McGuire,A., McKernan,K., McPheters,R.,
Meidirm,U., Meneus,L., Minova,T., Miranda,C., Mleaga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Sudmanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trifillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (15-JUN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 28, 2000 this sequence version replaced gi:8567832.
All repeats were identified using RepeatMasker:
smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

TITLE JOURNAL COMMENT

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L10415
Center clone name: 173_C-19
----- Summary Statistics -----
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 158085 bases at least Q40
Consensus quality: 168416 bases at least Q30
Consensus quality: 172386 bases at least Q20
Insert size: 174550; sum-of-ontigs
Quality coverage: 3.5 in Q20 bases; sum-of-ontigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1157: contig of 1157 bp in length

```


RESULT 15
AC009636/c
LOCUS AC009636
DEFINITION Mus musculus clone RP24-149H13, WORKING DRAFT SEQUENCE, 20
ACCESSION AC009636
VERSION AC009636.1 GI:16948003
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILLTOP.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eulheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Mammalia; Eulheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Birren, B., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 155765)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gaidyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, D., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schenck, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, K., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, J., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
JOURNAL Direct Submission
COMMENT Submitted (16-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
TITLE Genome Center
JOURNAL Center: Whitehead Institute/ MIT Center for Genome Research
COMMENT Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L17403
Center clone name: 149_H13
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 149108 bases at least Q40
Consensus quality: 152022 bases at least Q20
Consensus quality: 153025 bases at least Q20
Insert size: 158000; agarose-fp
Insert size: 153865; sum-of-contigs
Quality coverage: 12.8 in Q20 bases; sum-of-contigs
Quality coverage: 13.1 in Q20 bases; sum-of-contigs
NOTE: This is a 'working draft' sequence. It currently consists of 20 contigs. The true order of the pieces is not known and their order in this sequence reflects an arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.
1 8420: contig of 8420 bp in length
* 8421 8520: gap of 100 bp
* 8521 9147: contig of 627 bp in length
* 9148 9247: gap of 100 bp
* 9248 9681: contig of 634 bp in length
* 9682 9981: gap of 100 bp
* 9982 10632: contig of 651 bp in length
* 10633 10732: gap of 100 bp
* 10733 11428: contig of 696 bp in length
* 11429 11528: gap of 100 bp
* 11529 12139: contig of 611 bp in length
* 12140 12239: gap of 100 bp
* 12240 12867: contig of 628 bp in length
* 12868 12967: gap of 100 bp
* 12968 14377: contig of 1410 bp in length
* 14378 14477: gap of 100 bp
* 14478 15555: contig of 1078 bp in length
* 15556 15655: gap of 100 bp
* 15656 16770: contig of 1115 bp in length
* 16771 16870: gap of 100 bp
* 16871 17550: contig of 680 bp in length
* 17551 17650: gap of 100 bp
* 17651 20264: contig of 2614 bp in length
* 20265 20364: gap of 100 bp
* 20365 22093: contig of 1729 bp in length
* 22094 22193: gap of 100 bp
* 22194 23381: contig of 1188 bp in length
* 23382 23481: gap of 100 bp
* 23482 72883: contig of 49402 bp in length
* 72884 72983: gap of 100 bp
* 72984 85260: contig of 12277 bp in length
* 85261 85360: gap of 100 bp
* 85361 103628: contig of 18268 bp in length
* 103629 103728: gap of 100 bp
* 103729 124191: contig of 20463 bp in length
* 124192 124291: gap of 100 bp
* 124292 148926: contig of 24635 bp in length
* 148927 149026: gap of 100 bp
* 149027 155765: contig of 6739 bp in length.
FEATURES
source
1..155765
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-149H13"
/clone_lib="RPCI-24 Male Mouse BAC"
1..8420
/note="assembly_fragment
clone_end:SP6
vector_side:left"
misc_feature
8521..9147
/note="assembly_fragment"
9248..9681
/note="assembly_fragment"
9682..10632
/note="assembly_fragment"
10733..11428
/note="assembly_fragment"
11529..12139
/note="assembly_fragment"
12240..12867
/note="assembly_fragment"
12968..14377
/note="assembly_fragment"
14478..15555
/note="assembly_fragment"
15656..16770
/note="assembly_fragment"
16871..17550
/note="assembly_fragment"
17651..20264
/note="assembly_fragment"
misc_feature

```
misc_feature      20365..22093      /note="assembly_fragment"  
misc_feature      22194..23381      /note="assembly_fragment"  
misc_feature      23482..72883      /note="assembly_fragment"  
misc_feature      72984..85260      /note="assembly_fragment"  
misc_feature      85361..103628     /note="assembly_fragment"  
misc_feature      103729..124191     /note="assembly_fragment"  
misc_feature      124292..148926     /note="assembly_fragment"  
misc_feature      149027..155765     /note="assembly_fragment"  
misc_feature      clone_end:T7  
vector:side:right"
```

```
BASE COUNT      43072 a 33927 c 33544 g 43321 t 1901 others  
ORIGIN
```

```
Query Match      13.0%; Score 42.8; DB 2; Length 155765;  
Best Local Similarity 54.4%; Pred. No. 0.15; 72; Indels 0; Gaps 0;
```

```
Matches 86; Conservative 0; Mismatches 0; Gaps 0;
```

```
OY 31 atcatcaaacctcgcgatactcgaacactatactggaacaacaacatctacatcgtt 90  
      ||||| ||| ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| |||||  
DB 55553 ATCATCACCATTATCATCATCACCACCATTCATCACCATTCATCATCATTCATC 55494  
OY 91 ggtcagtgcaaaagtggttaacacttcatactcctctctgtactactgtttaagctatc 150  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 55493 ATCATCATTCACCATCATCATCATCATCATCATCATCATCATCATCATCATCATC 55434  
OY 151 tgcactggtgttatacaacatgaagttctgtactactac 188  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 55433 ATCACCATCACCATCACCACCATTCCTTTATATACAC 55396
```

```
Search completed: August 15, 2002, 09:47:14  
Job time: 10808 sec
```

. . . .